

Up from Limited Dialog Systems!

Giuseppe Riccardi

University of Trento
via Sommarive, 14
38050, Trento, Italy

riccardi@disi.unitn.it

Philipp Cimiano

Bielefeld University
Universitätsstraße 21–23
33615, Bielefeld, Germany

cimiano@cit-ec.uni-bielefeld.de

Alexandros Potamianos

Technical University of Crete
73100, Chania
Crete, Greece

potam@telecom.tuc.gr

Christina Unger

Bielefeld University
Universitätsstraße 21–23
33615, Bielefeld, Germany

cunger@cit-ec.uni-bielefeld.de

Abstract

In the last two decades, information-seeking spoken dialog systems (SDS) have moved from research prototypes to real-life commercial applications. Still, dialog systems are limited by the scale, complexity of the task and coverage of knowledge required by problem-solving machines or mobile personal assistants. Future spoken interaction are required to be multilingual, understand and act on large scale knowledge bases in all its forms (from structured to unstructured). The Web research community have striven to build large scale and open multilingual resources (e.g. Wikipedia) and knowledge bases (e.g. Yago). We argue that a) it is crucial to leverage this massive amount of Web lightly structured knowledge and b) the scale issue can be addressed collaboratively and design open standards to make tools and resources available to the whole speech and language community.

1 Introduction

In the last two decades, interactive spoken dialog systems (SDS) have moved from research prototypes to real-life commercial applications (Tur and De Mori, 2011). Generally, SDS are built for a specific task (e.g. call routing) with ad-hoc limited knowledge base and for a predefined target language. However, one major limitation in commercial SDS prototyping is that they are not easily and quickly extensible and portable to new domains or languages. Such porting requirements range from defining (or extending) a domain ontology to hand-crafting a new grammar or training stochastic mod-

els for speech recognition and understanding. These are the research and engineering goals motivating the PortDial project whose objectives include the engagement of the whole technical community. In the PortDial project we would like to engage researchers in building resources that may be generated via top-down processes (grammars), bottom-up processes (statistical models) or via a fusion of both. In this position paper we want to address the critical limitations of SDS systems: a) poor ability to cover the knowledge space and its interface to the SDS components (speech recognition, language understanding and dialog manager) and b) collaboratively design open standards to make tools and resources available to the whole speech and language community.

2 Exploiting top-down knowledge

There are at least three main kinds of structured knowledge sources that SDS modules may exploit: ontologies, grammars, and lexica. *Ontologies* explicitly model background knowledge about a certain domain. In the last years, many free and open collaboratively created resources have emerged, including large multi-lingual corpora such as Wikipedia, and broad-coverage ontologies, e.g. as part of the Linked Data Cloud (Bizer et al., 2009), either created manually or extracted automatically from existing data (such as DBpedia or Yago). However, while also *lexica* such as Wiktionary are available today on the Web, ontologies typically lack information about linguistic realization. For this reason, ontologies available on the Web are not directly exploitable by dialog systems. Linguistic in-

formation is commonly captured in *grammars*, that are either hand-crafted or created by means of machine learning techniques. In order to be able to generate high-quality grammars with as little manual effort as possible, we aim at (semi) automating the knowledge-based generation of lexica and grammars. To achieve this, it is crucial to leverage Web resources for enriching ontologies with lexical and linguistic information, i.e. information about how ontological concepts are lexicalized in different languages, capturing in particular lexical and syntactic variation (Unger et al., 2010). This knowledge-centered grammar generation process may be merged with methods for automatically inferring structure from lightly annotated corpus, including data harvested from the Web, in a bottom-up fashion (Tur and De Mori, 2011). For a dialog system to be able to exploit ontologies, lexica and grammars, these three resources need to be tightly aligned, i.e. they need to share domain-relevant vocabulary. For this alignment, we propose to build on Semantic Web standards, mainly in order to support the incorporation of already existing data, to share resources for SDS engineering, and facilitate collaborative knowledge engineering. From a larger perspective, such an approach has the potential of creating SDS resources (ontologies, lexica and grammars) that are strongly aligned with each other as well as with other resources available on the Web, thus fostering the creation of an eco-system of linked resources that can be reused to facilitate the process of engineering and porting a dialog system to new domains and languages.

3 Collaboratively building and sharing knowledge

Today the lack of reusable linguistic resources and annotated data hinders the rapid development of spoken dialog systems in industry and academia alike. Despite progress in standardization of the format of SDS grammars and semantic representations, the data proper has to be hand-crafted for new applications and languages with little or no automation available. We argue above that language engineering technology is now mature to help create such linguistic resources automatically or semi-automatically using data that is either harvested

from the web or via community crowdsourcing using the “collective wisdom of expert crowds”. Although providing linguistic resources and tools for cost-effective SDS development is important and relevant, a data pool that is not openly sharable and continuously enriched fails its purpose. It is thus important to guarantee the sustainability of the linguistic SDS resources engineered via a community that both uses and actively develops the data pool. Towards this end, we envision both a free and premium data exchange targeting non-commercial users that can maintain and enrich the free version of the data pool, and commercial speech services developers that can contribute to the premium data pool via an electronic marketplace. This is the model we are launching within the EC-funded PortDial project and aiming at involving the research community at large and existing communities for sharing linguistic resources, such as METANET and METASHARE¹. We believe that the creation of sharable SDS data and linguistic resources for both academic and commercial use will lead to the democratization of spoken dialog systems development, reduce the barrier to entry for new developers, as well as lead to improved technologies for authoring speech services.

References

- C. Bizer, T. Heath, and T. Berners-Lee. 2009. Linked data—the story so far. *International Journal on Semantic Web and Information Systems*, 14:9.
- G. Tur and R. De Mori, editors. 2011. *Spoken Language Understanding: Systems for Extracting Semantic Information from Speech*. Wiley.
- Christina Unger, Felix Hieber, and Philipp Cimiano. 2010. Generating LTAG grammars from a lexicon-ontology interface. In Srinivas Bangalore, Robert Frank, and Maribel Romero, editors, *Proceedings of the 10th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+10)*, pages 61–68, 06/2010.

¹<http://www.meta-net.eu/meta-share>